

1
2
3
4
5
6
7
8
9
10
11
12

Conserving Phylogenetic Diversity can be a Poor Strategy for Conserving Functional Diversity

Florent Mazel^{1,*}, Arne Mooers¹, Giulio Valentino Dalla Riva², and Matthew W. Pennell²

¹*Department of Biological Sciences, Simon Fraser University, Burnaby, BC V5A 1S6,,
Canada*

²*Department of Zoology and Biodiversity Research Centre, University of British
Columbia, Vancouver, BC V6T 1Z4, Canada;*

* Corresponding Author; flo.mazel@gmail.com

13 **Abstract**

14 For decades, academic biologists have advocated for making conservation decisions in
15 light of evolutionary history. Specifically, they suggest that policymakers should
16 prioritize conserving phylogenetically diverse assemblages. The most prominent
17 argument is that conserving phylogenetic diversity (PD) will also conserve diversity in
18 traits and features (functional diversity; FD), which may be valuable for a number of
19 reasons. The claim that PD-maximized ('maxPD') sets of taxa will also have high FD is
20 often taken at face value and in cases where researchers have actually tested it, they have
21 done so by measuring the phylogenetic signal in ecologically important functional traits.
22 The rationale is that if traits closely mirror phylogeny, then saving the maxPD set of taxa
23 will tend to maximize FD and if traits do not have phylogenetic structure, then saving the
24 maxPD set of taxa will be no better at capturing FD than criteria that ignore PD. Here, we
25 suggest that measuring the phylogenetic signal in traits is uninformative for evaluating
26 the effectiveness of using PD in conservation. We evolve traits under several different
27 models and, for the first time, directly compare the FD of a set of taxa that maximize PD
28 to the FD of a random set of the same size. Under many common models of trait
29 evolution and tree shapes, conserving the maxPD set of taxa will conserve more FD than
30 conserving a random set of the same size. However, this result cannot be generalized to
31 other classes of models. We find that under biologically plausible scenarios, using PD to
32 select species can actually lead to less FD compared to a random set. Critically, this can
33 occur even when there is phylogenetic signal in the traits. Predicting exactly when we
34 expect using PD to be a good strategy for conserving FD is challenging, as it depends on
35 complex interactions between tree shape and the assumptions of the evolutionary model.
36 Nonetheless, if our goal is to maintain trait diversity, the fact that conserving taxa based
37 on PD will not reliably conserve at least as much FD as choosing randomly raises serious
38 concerns about the general utility of PD in conservation.

39

40 **Keywords.** Conservation, Functional diversity, Evolutionary diversity, Trait evolution,
41 Species prioritization

42

43 In the face of the current biodiversity crisis, society needs to decide how to distribute
44 limited funds and effort to conservation. Conservation biologists and policy makers have
45 presented many proposals for making rational and scientific decisions about which
46 species warrant the most protection (Bottrill et al. 2008).

47 One prominent prioritization scheme uses evolutionary history to place a
48 quantitative value on species and sets of species. The idea is that when making
49 conservation policy, we should try to conserve the set of species or habitats that harbour
50 the greatest amount of evolutionary history (Vane-Wright et al. 1991). While there are
51 many, overlapping metrics for measuring the evolutionary history encompassed by a set
52 of species (Winter et al. 2013; Tucker et al. 2016), the most common is the sum of all
53 branch lengths connecting a set of species to a common root (Faith 1992), called
54 Phylogenetic Diversity (PD). This measure is vague insofar as the units of “branch
55 length” are unspecified, but it is the metric whose maximization has been proposed as a
56 conservation prioritization strategy.

57 While PD has been only sparingly used in actual policy decisions (for one
58 example, see the EDGE program of the Zoological Society of London;
59 www.edgeofexistence.org), it has caught the attention of researchers; according to
60 Google Scholar, the original Faith (1992) paper on the topic has been cited more than
61 1900 times as of April 2017. Indeed, Faith’s paper has spawned an entire subfield in
62 which biologists and mathematicians have worked out complex solutions to measuring
63 and maximizing PD (e.g. Rodrigues and Gaston 2002; Forest et al. 2007; Bordewich et al.
64 2008; Bennett et al. 2014; Chao et al. 2015; Pollock et al. 2015; Thuiller et al. 2015).
65 Faith and other researchers have proposed several key reasons why conserving PD is
66 worthwhile. Prioritizing species’ conservation to maximize PD may help ensure that: (i)
67 remarkable species that occur as evolutionary isolated lineages (e.g. tailed frogs, tuataras,
68 *Welwitschia*) are prioritized (Rosauer and Mooers 2013); (ii) essential ecosystem
69 functions and services are maintained (Cadotte et al. 2008, but see Srivastava and
70 Vellend, 2005 for a discussion on the link with applied conservation); and that (iii) we
71 maximize ‘evolutionary potential’ (Faith 1992; Forest et al. 2007). All of these ideas have
72 are underpinned by the claim that phylogenetically diverse sets of taxa contain a
73 disproportionately large amount of trait/feature/functional diversity. Hereafter, we will

74 not make a distinction between trait, feature and functional diversity and we will refer to
75 them as functional diversity, or FD.

76 Like evolutionary history, functional diversity is an ambiguous concept with
77 many potential measures. Villéger et al. (2008) suggest that FD has three components:
78 richness, divergence, and evenness. Functional richness generally measures ‘how much
79 trait space is filled, while functional divergence and evenness indices describe how this
80 space is filled’ (Schleuter et al. 2010). Functional richness represents the amount of
81 functional trait space that is encapsulated by a set of species, is usually correlated with
82 species richness, and can be related to the functioning of ecosystems (see, e.g. Cadotte et
83 al. 2011). The second component, functional divergence, is largely independent of
84 species richness and describes how species are clustered in trait space, which may be
85 valuable to conservation biologists interested in, e.g., ecosystem services (e.g. Díaz et al.
86 2007). These two classes of measure are often used in trait-ecology and conservation (e.g.
87 Devictor et al. 2010; Mouillot et al. 2014) and we assume here that conserving functional
88 richness and/or divergence is a valuable conservation objective. While functional
89 divergence relates to some measure of mean trait distances between species, functional
90 evenness relates to the variance of these trait distances. This last FD dimension describes
91 the extent to which species are clustered with their (direct) neighbours versus being
92 regularly spaced in trait space. We did not consider any measure of functional evenness
93 (such as the Functional Evenness Index, Villéger et al. 2008) in what follows because
94 could not identify any potential causal link between evenness in trait space and
95 ecosystem function or services, and trait evenness has generally not been a concern of
96 conservation biologists.

97 In this paper, we ask whether maximizing PD help to conserve functional
98 diversity. The common rationale for using PD as a proxy for FD is that many ecologically
99 relevant traits harbour some degree of phylogenetic signal (see, e.g., Winter et al. 2013).
100 At a glance, this seems logical: if the data shows strong phylogenetic signal, then picking
101 distantly related taxa seems a sensible way to ensure that you have captured species from
102 across trait space. And indeed, if we assume that traits have evolved according to a
103 Brownian motion (BM, Felsenstein 1985) process, then this will be true (see below). The
104 converse is also true: if traits do not show phylogenetic signal, other methods for

105 capturing FD are needed (see, e.g., Faith 2015). A number of studies from across
106 evolutionary biology, ecology, and conservation biology have evaluated the amount of
107 phylogenetic signal (measured in a variety of ways, see Münkemüller et al. 2012) in
108 ecologically important traits (see, e.g., Freckleton et al. 2002; Blomberg et al. 2003;
109 Chamberlain et al. 2012). Recently, Kelly et al. (2014) specifically focused on the
110 implications of phylogenetic signal for the use of PD in conservation. They constructed
111 trees using a wide variety of morphological traits and found that while closely related
112 species often shared many trait combinations, these traits were not informative for deeper
113 splits in the tree. They argued that this was evidence that maximizing PD would not
114 reliably maximize feature diversity.

115 The results of these studies (along with, likely many more) have been widely
116 variable: some traits in some taxa in some regions contain a lot of phylogenetic signal
117 while others do not. This led Winter et al. (2013, p 201) to conclude: “If the conservation
118 goal is to conserve functional diversity, considering phylogenetic diversity might be
119 either well suited or totally misleading”. We argue that there is an important and
120 underappreciated assumption in this line of reasoning: that the degree of phylogenetic
121 signal in some key trait(s) is indicative of the effectiveness of using PD to conserve FD.

122 There are two reasons to be suspicious of this assumption. First, our thinking
123 about phylogenetic signal has been informed by considering a few simple models of trait
124 evolution; other, completely different, classes of models may generate variation in
125 phylogenetic signal that are far less intuitive. Second, the motivating idea is that policy
126 makers should use PD to pick sets of taxa to conserve. These sets are, by definition, non-
127 random and therefore may have different statistical properties from the clade as a whole.
128 In this paper, we simulate data under different models of evolution and, for the first time
129 to our knowledge, directly test how much FD the set of taxa that maximize PD (‘the
130 maxPD set’ hereafter) contains compared to alternative possible sets.

131 Specifically, we contrast the outcome for FD conservation of conserving the
132 maxPD set of taxa -- and letting everything else go extinct -- with conserving a random
133 set of taxa of the same size. Here, random simply means conservation decisions that
134 ignore phylogenetic position and the functional traits we are considering. As such,
135 random sets provide a natural point of comparison to understand the properties of

136 maxPD. We note that we are not testing whether conserving maxPD will maximize the
137 amount of FD it is possible to conserve. While this claim is likely what some advocates
138 of PD have in mind, and is what Kelly et al. (2014) actually aimed to test, it is a rather
139 high bar to meet. Indeed, it is easy to concoct scenarios in which this will not hold; if, for
140 example, traits were so labile that there was no phylogenetic signal (i.e., the “white
141 noise” model), then we would expect that maxPD sets would contain no more or less FD
142 on average than any other set. It therefore seems too high a bar to expect for PD to
143 *always* maximize FD in order to declare it useful for conserving FD. Instead, we believe
144 we must first clear a much lower bar -- does prioritizing species based on maximizing PD
145 do better at capturing FD than prioritizing a random set?

146 Below, we demonstrate that both the model of trait evolution and the tree shape
147 are relevant for deciding whether or not PD is a good strategy for conserving FD. And,
148 more surprisingly, we show that even when there is phylogenetic signal in the sampled
149 traits, using PD to guide conservation decisions can lead to choice outcomes for
150 conserving FD that are worse than if we were choosing randomly. This counter-intuitive
151 result suggests that we need to re-assess both the way in which we intuitively consider
152 phylogenetic signal in conservation biology, and the justification for phylogenetically-
153 based prioritization.

154

155 **METHODS**

156

157 We wanted to test the following conjecture under a variety of evolutionary
158 scenarios:

159 If we select a set S of m taxa from a clade of size n such that the sum of the
160 branch lengths connecting S is at least as large as that stemming from any other possible
161 subset (i.e. PD is maximized), then S will contain at least as much FD on average as a
162 randomly chosen subset of size m .

163 Four things are notable about this test. First, as stated above, we are not trying to
164 determine whether the maxPD set will actually maximize FD (i.e., that S would contain at
165 least as much FD as any other set of the same size). Second, we are interested in the

166 expectation, or average. Evolution certainly can take interesting turns such that some sub-
167 clades span the functional diversity of the entire group (e.g., different clades of African
168 rift cichlids have independently evolved the same breadth of functional diversity in
169 different lakes; Muschick et al. 2012). Or, a trait important for ecosystem functioning
170 may also evolve only once and we would like to make sure we capture this lineage
171 (Davies et al. 2016). Average properties are critical, however, because PD's utility in
172 conservation comes precisely when we don't know the traits or functions that matter; the
173 best we can hope for is that, on average, we expect it to perform well. Third, we do not
174 require S to uniquely maximize PD. We use the greedy algorithm proposed by Bordewich
175 et al. (2008) to find our maxPD set of species S . For a given tree there are likely multiple,
176 and possibly very many, sets of with the same PD as S . As this number will vary across
177 simulations and could, in some case, be very large, we have chosen to select only one set
178 per simulation. This allows us to carry out more simulations, increasing the generality of
179 our results. And last, we are assuming that all of the taxa we select will survive and that
180 every other taxa in the clade will go extinct with certainty. This is, of course,
181 unreasonable and unrealistic but is useful for the purposes of illustration (see Discussion).

182

183 *Simulations*

184

185 To explore a broad range of tree shapes, we simulated trees under three different
186 diversification models. First, we simulated trees under a Yule process (no extinction).
187 Second, to obtain trees that were more 'tippy' (i.e., having more speciation events close
188 to the present), we used a coalescent model. In both cases, we simulated trees with 32 and
189 with 64 taxa. To obtain trees that were more unbalanced than those typically produced by
190 the Yule or coalescent processes, we simulated trees where the speciation rate evolved as
191 a continuous trait along the tree (Rabosky 2010; Beaulieu and O'Meara 2015). This
192 allowed some groups within a tree to diversify faster than others, with this heterogeneity
193 being phylogenetically clustered.

194 To do the latter, we used R scripts from Beaulieu (2015, modified from Rabosky
195 2010) and set the initial speciation rate to .06. Each tree was subsequently pruned to $n =$
196 64 and to $n = 32$. We then kept the 100 first trees that encompassed a wide range of

197 imbalance values: we kept 10 trees by bins of 0.4 imbalance value (as measured by β ,
198 Blum et al. 2006) from -1.6 to 2. For a point of comparison, we also used fully
199 imbalanced ($\beta = -2$) and balanced trees ($\beta = 10$) of 32 and 64 species.

200 To explore a range of continuous trait evolution models, we used 1) the BM
201 model setting the drift parameter $\sigma^2 = 1$ (we did not explore multiple values of σ^2 because
202 it does not influence the phylogenetic signal of the data and thus will not impact our
203 results); 2) the Ornstein-Uhlenbeck (OU, Hansen 1997), with $\sigma^2 = 1$ and $\alpha = \{1.4, 7\}$
204 corresponding to half life of .1 and .5 for a tree with total height rescaled to 1; and 3) the
205 early burst (EB, Harmon et al. 2010, $r = -5$ and -1). For discrete traits, we used the
206 Markov model of evolution (Pagel 1994). We used a simple Markov model with 4
207 character states and all transitions rates equal to .1 or 1. Speciation models, in which
208 trait evolution occurs (at least in part) when lineages split, were also used for both
209 continuous and discrete traits by applying a Pagel κ transformation to the original tree
210 (Pagel 1999). We simulated datasets with $N = \{1, 2, 4\}$ independently evolving traits. As
211 we wanted to keep the simulations simple, we did not include variations such as multi-
212 rate BM (O’Meara et al. 2006; Eastman et al. 2011) or multi-optima OU models (Butler
213 and Kings 2004; Ingram and Mahler 2013; Uyeda and Harmon 2014). For each set of
214 parameters and number of trait values, we simulated 1000 datasets. In each case, we also
215 computed the phylogenetic signal contained in the data by calculating the spearman
216 correlation between phylogenetic and trait distance matrices (following Kelly et al.,
217 2014) in addition to more commonly used measures such as Blomberg’s K (Blomberg et
218 al. 2003); Blomberg’ K does not allow for the possibility of “anti-signal” in traits,
219 wherein close relatives are more dissimilar than distantly related taxa.

220

221 *Analysis*

222

223 For each dataset, we selected two sets of m species ($m = \{8, 16\}$) out of the total
224 number of n species in the tree ($n = \{32, 64\}$) for a total of four parameter combinations.
225 One set was chosen at random and the other was a set that maximized PD (i.e., maxPD),
226 using the algorithm of Bordewich et al. 2008 that we implemented in R. We then
227 computed FD for both the random and the maxPD sets. Functional richness was

228 estimated using the convex hull volume (Cornwell et al. 2006), which measures the total
229 volume encapsulated by all species in trait space. In a single dimension, this simply
230 equals the range of values. This broadly used metric in ecology is set monotonic with
231 species richness, a property generally assumed desirable in conservation whereby the
232 addition of a new species can never decrease the metric's value (Ricotta 2005). Functional
233 divergence was estimated using Rao's quadratic entropy (Rao 1982; Botta-Dukát 2005),
234 which represents the mean trait distance between pairs of species (including the null
235 distance of a species with itself) and is highly correlated to the trait variance across tips
236 (de Bello et al. 2016). While this index is not set monotonic with species richness, we
237 feel that it might be of interest to test the robustness of our results. By using functional
238 richness and functional divergence, we are able to capture both the spread of the data in
239 trait space as well as how clustered it is, since it is not immediately clear what quantity is
240 most relevant for the use of PD in conservation. For discrete traits, the convex hull
241 volume is less meaningful than for continuous traits. Therefore, we used the number of
242 unique trait states in the set as a measure of FD for discrete traits (Petchey and Gaston
243 2006; Mouillot et al. 2014).

244 For each simulation, we then computed the relative amount of FD in the two sets
245 using the following metric:

$$246 \quad rFD = FD_{\max PD} / (FD_{\text{Random}} + FD_{\max PD}) \quad (\text{eq. 1})$$

248
249 A rFD value of greater than or equal to 0.5 means that the PD set contains at least
250 as much FD as the random set and a rFD value less than 0.5 means that it contains less.
251 All analysis were run in R, making special use of the ape (Popescu et al. 2012), ade4
252 (Dray et al. 2007), phytools (Revell 2012), geiger (Pennell et al. 2014a), geometry (Habel
253 et al. 2015), apTreeshape (Bortolussi et al. 2006), and mvMORPH (Clavel et al. 2015)
254 packages. All code to run the analyses is available at
255 https://github.com/FloMazel/PD_FD.

256

257 **RESULTS**

258 We found that, under many common models of trait evolution, conserving the maxPD set
259 of taxa will on average conserve more FD than conserving a random set of the same size
260 (i.e. rFD is always $> .5$, see Table 1, note that rFD is an average over all simulations but
261 individual simulation may have $rFD < .5$). This is because related species tend to be on
262 average closer in trait space than distantly ones (Figure 2a-d), so that selecting distantly
263 related species increases FD. This result is more pronounced for very early evolution (as
264 modelled by an early burst model of evolution) because in this case distantly related
265 species are always well separated in the functional space. On the contrary, very late
266 evolution, or very strong stabilizing selection (as modelled by the OU process) tends to
267 erase the differences between set of species, but never leads (on average) to the maxPD
268 set of species to harbour less FD than the random set. Overall, an increase of
269 phylogenetic signal tends to increase the difference between FD of the two sets (Table 1
270 and Supp. Tables). Our results also hold for alternative tree sizes (Supp. Table 1) and
271 Functional Divergence (measured as Rao's Quadratic entropy, see Supp. Table 2). Also,
272 the difference between FD of the two sets of species is largest when a small proportion of
273 tree size is selected and tends to decrease when more species are selected (Supp. Table
274 2). This is expected: if 100% of the species are selected, the FD of the random and
275 maxPD sets will be equal and equal to the FD of the entire clade ($rFD = .5$).

276 However, this result cannot be generalized to all classes of models. When traits
277 evolve on an imbalanced tree under a speciation model (Figure 1), early diverging
278 species are always selected to maximize PD (species No 1, 2 and 3 in Figure 2d-g) but
279 are functionally relatively similar since their traits have not diverged much. Here, a
280 random choice of species will, on average, select species that are much less functionally
281 similar, yielding higher FD and thus an $rFD < .5$ (Figure 2d-h). As with other models, the
282 difference between FD of the two sets of species is strongest when a small proportion of
283 tree size is selected and tends to decrease when more species are selected (Fig. S1). This
284 result also holds using Rao's measure of Functional Divergence (Fig. S2).

285 Above we described the results for $n=2$ traits. Multiple traits are likely important
286 for maintaining ecosystem functions and services and for potentially promoting
287 diversification. However, our results do not qualitatively depend on how many traits we
288 consider. If we use convex hull volume as a measure of FD, then the patterns we see in

289 one or two dimensions are only exacerbated in higher trait dimensions (Fig. S1): in cases
290 where maxPD does poorly, adding more traits makes it do worse, and in cases where it
291 does well, more traits accentuate its success. When we measured FD using Rao's
292 quadratic entropy, there was no difference between results at two or higher dimensions
293 (Fig. S2). This is because Rao's quadratic entropy represents the mean functional
294 distance between species (including comparing a species to itself) and we know that, for
295 a BM model, increasing the number of traits simply decreases the variance of functional
296 distances between species (see e.g. Letten and Cornwell 2015) and thus will not impact
297 the average of the rFD metric. Importantly all our results are also robust to variation in
298 tree size and number of selected species (Fig. S1-2) and also hold when a speciation
299 model of evolution for discrete traits is applied instead (i.e. a Markov model, see Fig. S3).

300 After seeing our results, we naively thought that if there was a non-negative
301 correlation between the traits and the phylogeny (i.e., "phylogenetic signal" broadly
302 construed), this would mean that PD should on average do at least as well as random. Our
303 intuition here was wrong. Indeed, even in our "worst case" scenario, when the tree is
304 perfectly imbalanced and trait evolution only occurs at speciation, the correlation
305 between the trait covariance matrix and the phylogenetic covariance matrix is still
306 positive -- close relatives resemble one another but selecting the maxPD set of taxa
307 captures less FD than a randomly chosen set on average (Fig S4-5)! The key to resolving
308 this apparent paradox is recognizing that the phylogenetic signal of the entire dataset is
309 not expected to equal the phylogenetic signal of non-random subsets of the data. In
310 particular, the set of species that maximized PD is expected to occupy a very particular
311 position in the phylogenetic and functional distances space.

312 To intuitively understand this point, we present in figure 2 (panels d-h) a
313 simplified toy example with a fully imbalanced phylogeny of 16 species from which four
314 species are selected, either at random (squares in the figure) or in order to maximize PD
315 (maxPD set, represented by triangles). In this case, species 1, 2 and 3 will always be
316 selected to maximize PD, while the fourth one will be chosen at random among the
317 remaining species (Fig 2d). In the case of a speciation model of evolution, three out of
318 four species from the maxPD set (species 1,2 and 3) will be, on average, relatively
319 clumped in the trait space (Fig. 2e, triangles) and thus harbour small trait distances, while

320 being distantly related in the phylogeny (Fig. 2g). On the contrary, the random subset
321 (squares) will be more spread in the trait space (Fig. 2f) and thus harbour relatively
322 higher trait distances, while being relatively less distant in the phylogeny (Fig. 2g). So,
323 the random set will harbour more FD than the maxPD set (Fig 2h). While the overall (i.e.
324 for all species) relationship between trait and phylogenetic distances is slightly positive
325 (and not negative), the same relationship restricted to random and maxPD subsets
326 becomes negative (imagine a line between squares and triangles on figure 2g). It thus
327 appears that the overall trend between all species are not representative of the trend
328 between members of the maxPD and random sets; the measure of phylogenetic signal on
329 the whole phylogeny may not be a good proxy for the representativeness of FD by the
330 maxPD set of species.

331

332 **DISCUSSION**

333 Most of the arguments for using PD in conservation decisions reason that
334 conserving phylogenetically diverse sets of taxa is valuable because it conserves some
335 sort of trait diversity; for other rationales for conserving PD see e.g., Vane-Wright et al.
336 (1991) or Rosauer and Mooers (2013). Trait diversity may be valuable if it helps maintain
337 ecosystem functioning and services (e.g. Best et al. 2013; Winter et al. 2013; Gross et al.
338 2017) , if it captures ‘evolutionary potential’ (Faith 1992; Forest et al. 2007), or if trait
339 diversity increases the probability of encompassing rare traits that are deemed valuable
340 for their rareness *per se* (e.g. egg-laying in mammals, Rosauer and Mooers, 2013). Here,
341 we are agnostic as to why traits are valuable to conserve; we only assume that they are.

342 Our main results speak to at least on other recent paper that also purported to test
343 whether PD was a good proxy of feature diversity. Using a wide variety of morphological
344 traits previously used to infer phylogenies, Kelly et al. (2014) showed that, while closely
345 related species often share many trait combinations, these traits are not informative for
346 deeper splits in the tree – i.e. that phylogenetic signal decays rapidly in the tested
347 character matrices. A second key finding of the Kelly et al. study was that the trait
348 distances between the two most distant species in the tree (i.e. considering FD of the
349 maxPD sets of two species) is lower than the maximal trait distance in the dataset. Our
350 test is both more stringent and more general than that of Kelly et al. First, we did not test

351 whether preserving the maxPD species will maximize the amount of FD it is possible to
352 preserve, but rather if the maxPD set capture more FD than a random set, a much lower
353 bar to meet. For example, even in the situation where we found the maxPD set to harbour
354 more FD than random (e.g. in the case of a simple BM model), it is likely that this set
355 does not maximize FD. Second, while Kelly et al. focused on the FD of the maxPD set
356 that comprises only two species, we consider here sets of taxa with a broader range of
357 sizes (8 and 16 species). This allowed us to show that the measure of phylogenetic signal
358 on the whole phylogeny may not be a good proxy for the representativeness of FD by the
359 maxPD set of species.

360 Our analysis is, of course, rather oversimplified in some ways. In the real world,
361 we do not have full control over which species survive and which are lost. Conservation
362 prioritization itself is a result of a complex interplay of social, economic, political, and
363 scientific priorities and is not always species-centred. And even if we did have the power
364 to decide, we would neither conserve everything we chose, nor would everything we
365 didn't choose go extinct. Furthermore, the extinction proportions used in our simulations
366 (e.g. 75%) are beyond dystopic. But the simplicity of our simulations allows us to
367 evaluate the logic underlying the (seemingly obvious, but not actually obvious at all)
368 claim that conserving phylogenetic diversity will result in conserving trait diversity. We
369 realize also that some of the situations which produce rFD values of less than 0.5 may not
370 be biologically realistic. It is unlikely that *most* trait evolution is speciational (Pennell et
371 al. 2014b) and, while empirical trees are more unbalanced than those produced by Yule
372 models (Mooers and Heard 1997), totally unbalanced trees are rare. While, such extreme
373 scenarios are not necessary to reliably get rFD values of less than 0.5, we think that these
374 cases are useful for critically evaluating the underlying logic behind the use of PD and
375 will perhaps stimulate the production of more direct tests of the usefulness of PD to
376 represent FD.

377 While there have been several meta-analyses comparing the fit of various trait
378 models across clades (Harmon et al. 2010; Pennell et al. 2015), these have been limited to
379 a few simple models, all of which are in the part of parameter space where PD performs
380 well as a proxy for FD. More comprehensive meta-analyses of the fit of models to
381 comparative data are required to allow us to assess where in model space traits of interest

382 generally fall. Furthermore, recent innovations using simulation-based approaches (e.g.
383 Slater et al. 2012; Sukumaran et al. 2016; Clarke et al. 2017) may allow us to expand
384 beyond our limited set of process models. A simpler empirical test of the utility of PD is
385 to gather empirical datasets and to repeat our analytical procedure on these. We would
386 then be able to ask for these empirical datasets whether the maxPD set of taxa will
387 contain more FD than a randomly chosen set. To our knowledge, no such test has been
388 performed. While this test would not provide a definitive answer to the utility of PD, it
389 would at least provide some indication of how concerned we should be given our results.

390 That said, if we had some approximate idea as to how likely it is the maxPD fails
391 to capture FD, policy recommendations might still be difficult. If maxPD does better than
392 random in, say, 80% of clades/traits, should this be interpreted as an endorsement of the
393 use of PD in conservation or a denouncement? What level of increase in FD is important?
394 A formal decision-theoretic framework (Robert 2007) might be needed for navigating
395 these thorny problems.

396

397 **CONCLUSION**

398 Given the interest in using PD in conservation decisions and the amount of work
399 that has gone into the problem of how to measure and prioritize PD, it is surprising that
400 there has not been direct theoretical or broad empirical evaluations of what exactly PD
401 captures. Here, we find that under many common models of trait evolution and tree
402 shapes, conserving the maxPD set of taxa will indeed conserve more FD than conserving
403 a random set of the same size. However, under other biologically plausible scenarios,
404 using PD to select species can actually lead to less FD compared to a random set.
405 Importantly, this can occur even when there is phylogenetic signal in the traits. The fact
406 that conserving taxa based on PD will not always reliably conserve at least as much FD
407 as choosing randomly may raise serious concerns about the utility of PD in conservation
408 if our goal is to save a diverse set of traits.

409

410

411 **Acknowledgements**

412 This paper is a joint effort of the sCAP working group funded by the Synthesis Centre of
413 the German Centre for Integrative Biodiversity Research Halle-Jena-Leipzig and the
414 Canadian Institute for Ecology and Evolution. FM was supported by an NSERC
415 Accelerator Grant to AOM. GDR thanks Mike Steel. We thank Risa Sergeant at the
416 University of Ottawa for hospitality, the Crawford Lab for Evolutionary Studies at SFU
417 for discussion, Luke Harmon and Will Cornwell for statistical advice, Caroline Tucker,
418 William Pearse, Marc Cadotte, Wilfried Thuiller, Laura Pollock, Thomas Buckley, and
419 two anonymous reviewers for comments on a previous version of this manuscript.

420 **References**

- 421 Beaulieu J.M., O'Meara B.C. 2015. Extinction can be estimated from moderately sized
422 molecular phylogenies. *Evolution* (N. Y). 69:1036–1043.
- 423 de Bello F., Carmona C.P., Lepš J., Szava-Kovats R., Pärtel M. 2016. Functional
424 diversity through the mean trait dissimilarity: resolving shortcomings with existing
425 paradigms and algorithms. *Oecologia*. 180:933–940.
- 426 Bennett J.R., Elliott G., Mellish B., Joseph L.N., Tulloch A.I.T., Probert W.J.M., Di
427 Fonzo M.M.I., Monks J.M., Possingham H.P., Maloney R. 2014. Balancing
428 phylogenetic diversity and species numbers in conservation prioritization, using a
429 case study of threatened species in New Zealand. *Biol. Conserv.* 174:47–54.
- 430 Best R.J., Caulk N.C., Stachowicz J.J. 2013. Trait vs. phylogenetic diversity as predictors
431 of competition and community composition in herbivorous marine amphipods. *Ecol.*
432 *Lett.* 16:72–80.
- 433 Blomberg S.P., Garland T.J.R., Ives A. 2003. Testing for phylogenetic signal in
434 comparative data : behavioral traits are more labile. *Evolution* (N. Y). 57:717–745.
- 435 Blum M., François O., Steel M. 2006. Which Random Processes Describe the Tree of
436 Life? A Large-Scale Study of Phylogenetic Tree Imbalance. *Syst. Biol.* 55:685–691.
- 437 Bordewich M., Rodrigo A., Semple C., Collins T. 2008. Selecting Taxa to Save or
438 Sequence: Desirable Criteria and a Greedy Solution. *Syst. Biol.* 57:825–834.
- 439 Bortolussi N., Durand E., Blum M., Francois O. 2006. apTreeshape: statistical analysis of
440 phylogenetic tree shape. *Bioinformatics*. 22:363–364.
- 441 Botta-Dukát Z. 2005. Rao' s quadratic entropy as a measure of functional diversity based

- 442 on multiple traits. *J. Veg. Sci.* 16:533–540.
- 443 Bottrill M.C., Joseph L.N., Carwardine J., Bode M., Cook C., Game E.T., Grantham H.,
444 Kark S., Linke S., McDonald-Madden E., Pressey R.L., Walker S., Wilson K.A.,
445 Possingham H.P. 2008. Is conservation triage just smart decision making? *Trends*
446 *Ecol. Evol.* 23:649–654.
- 447 Butler M.A., Kings A.A. 2004. Phylogenetic comparative analysis: A modeling approach
448 for adaptive evolution. *Am. Nat.* 164:683–695.
- 449 Cadotte M.W., Cardinale B.J., Oakley T.H. 2008. Evolutionary history and the effect of
450 biodiversity on plant productivity. *Proc. Natl. Acad. Sci. U. S. A.* 105:17012–7.
- 451 Cadotte M.W., Carscadden K., Mirotchnick N. 2011. Beyond species: functional
452 diversity and the maintenance of ecological processes and services. *J. Appl. Ecol.*
453 48:1079–1087.
- 454 Chamberlain S.A., Hovick S.M., Dibble C.J., Rasmussen N.L., Van Allen B.G., Maitner
455 B.S., Ahern J.R., Bell-Dereske L.P., Roy C.L., Meza-Lopez M., Carrillo J., Siemann
456 E., Lajeunesse M.J., Whitney K.D. 2012. Does phylogeny matter? Assessing the
457 impact of phylogenetic information in ecological meta-analysis. *Ecol. Lett.* 15:627–
458 636.
- 459 Chao A., Chiu C.-H., Hsieh T.C., Davis T., Nipperess D.A., Faith D.P. 2015. Rarefaction
460 and extrapolation of phylogenetic diversity. *Methods Ecol. Evol.* 6:380–388.
- 461 Clarke M., Thomas G.H., Freckleton R.P. 2017. Trait Evolution in Adaptive Radiations:
462 Modeling and Measuring Interspecific Competition on Phylogenies. *Am. Nat.*
463 189:121–137.
- 464 Clavel J., Escarguel G., Merceron G. 2015. mv morph : an r package for fitting
465 multivariate evolutionary models to morphometric data. *Methods Ecol. Evol.*
466 6:1311–1319.
- 467 Cornwell W.K., Schwilk D.W., Ackerly D.D., Schwilk L.D.W. 2006. A trait-based test
468 for habitat filtering: convex hull volume. *Ecology.* 87:1465–1471.
- 469 Davies T.J., Urban M.C., Rayfield B., Cadotte M.W., Peres-Neto P.R. 2016.
470 Deconstructing the relationships between phylogenetic diversity and ecology: a case
471 study on ecosystem functioning. *Ecology.* 97:2212–2222.
- 472 Devictor V., Mouillot D., Meynard C., Jiguet F., Thuiller W., Mouquet N. 2010. Spatial

473 mismatch and congruence between taxonomic, phylogenetic and functional
474 diversity: the need for integrative conservation strategies in a changing world. *Ecol.*
475 *Lett.* 13:1030–1040.

476 Díaz S., Lavorel S., de Bello F., Quétier F., Grigulis K., Robson T.M. 2007.
477 Incorporating plant functional diversity effects in ecosystem service assessments.
478 *Proc. Natl. Acad. Sci. U. S. A.* 104:20684–9.

479 Dray S., Dufour A.-B., Dray S., Dufour A.-B. 2007. The ade4 Package: Implementing the
480 Duality Diagram for Ecologists. *J. Stat. Softw.* 22.

481 Eastman J.M., Alfaro M.E., Joyce P., Hipp A.L., Harmon L.J. 2011. A novel comparative
482 method for identifying shifts in the rate of character evolution on trees. *Evolution.*
483 65:3578–89.

484 Faith D.P. 1992. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.*
485 61:1–10.

486 Faith D.P. 2015. The unimodal relationship between species' functional traits and habitat
487 gradients provides a family of indices supporting the conservation of functional trait
488 diversity. *Plant Ecol.* 216:725–740.

489 Felsenstein J. 1985. Phylogenies and the comparative method. *Am. Nat.*:1–15.

490 Forest F., Grenyer R., Rouget M., Davies T.J., Cowling R.M., Faith D.P., Balmford A.,
491 Manning J.C., Procheş S., van der Bank M., Reeves G., Hedderson T.A.J.,
492 Savolainen V. 2007. Preserving the evolutionary potential of floras in biodiversity
493 hotspots. *Nature.* 445:757–760.

494 Freckleton R.P., Harvey P.H., Pagel M. 2002. Phylogenetic analysis and comparative
495 data: a test and review of evidence. *Am. Nat.* 160:712–726.

496 Gross N., Bagousse-Pinguet Y. Le, Liancourt P., Berdugo M., Gotelli N.J., Maestre F.T.
497 2017. Functional trait diversity maximizes ecosystem multifunctionality. *Nat. Ecol.*
498 *Evol.* 1:132.

499 Habel K., Grasman R., Gramacy R.B., Stahel A., Sterratt D.C. 2015. geometry: Mesh
500 Generation and Surface Tessellation. R package version 0.3-6. Available from
501 <https://cran.r-project.org/package=geometry>.

502 Hansen T.F. 1997. Stabilizing Selection and the Comparative Analysis of Adaptation.
503 *Evolution (N. Y).* 51:1341.

- 504 Harmon L.J., Losos J.B., Jonathan Davies T., Gillespie R.G., Gittleman J.L., Bryan
505 Jennings W., Kozak K.H., McPeck M.A., Moreno-Roark F., Near T.J., Purvis A.,
506 Ricklefs R.E., Schluter D., Schulte II J.A., Seehausen O., Sidlauskas B.L., Torres-
507 Carvajal O., Weir J.T., Mooers A.Ø. 2010. Early bursts of body size and shape
508 evolution are rare in comparative data. *Evolution* (N. Y). 64:2385–2396.
- 509 Ingram T., Mahler D. 2013. SURFACE: detecting convergent evolution from
510 comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike
511 Information Criterion. *Methods Ecol. Evol.* 4:416–425.
- 512 Kelly S., Grenyer R., Scotland R.W. 2014. Phylogenetic trees do not reliably predict
513 feature diversity. *Divers. Distrib.* 20:600–612.
- 514 Letten A.D., Cornwell W.K. 2015. Trees, branches and (square) roots: why evolutionary
515 relatedness is not linearly related to functional distance. *Methods Ecol. Evol.* 6:439–
516 444.
- 517 Mooers A.O., Heard S.B. 1997. Inferring Evolutionary Process from Phylogenetic Tree
518 Shape. *Q. Rev. Biol.* 72:31–54.
- 519 Mouillot D., Vileger S., Parravicini V., Kulbicki M., Arias-Gonzalez J.E., Bender M.,
520 Chabanet P., Floeter S.R., Friedlander A., Vigliola L., Bellwood D.R. 2014.
521 Functional over-redundancy and high functional vulnerability in global fish faunas
522 on tropical reefs. *Proc. Natl. Acad. Sci.* 111:13757–13762.
- 523 Münkemüller T., Lavergne S., Bzeznik B., Dray S., Jombart T., Schiffrers K., Thuiller W.
524 2012. How to measure and test phylogenetic signal. *Methods Ecol. Evol.* 3:743–756.
- 525 Muschick M., Indermaur A., Salzburger W. 2012. Convergent evolution within an
526 adaptive radiation of cichlid fishes. *Curr. Biol.* 22:2362–8.
- 527 O’Meara B.C., Ané C., Sanderson M.J., Wainwright P.C. 2006. Testing for different rates
528 of continuous trait evolution using likelihood. *Evolution* (N. Y). 60:922–933.
- 529 Pagel M. 1994. Detecting correlated evolution on phylogenies: a general method for the
530 comparative analysis of discrete characters. *Proc. R. Soc. London, Ser. B.* 255:37–
531 45.
- 532 Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature.* 401:877–
533 884.
- 534 Pennell M.W., Eastman J.M., Slater G.J., Brown J.W., Uyeda J.C., FitzJohn R.G., Alfaro

- 535 M.E., Harmon L.J. 2014a. geiger v2.0: an expanded suite of methods for fitting
536 macroevolutionary models to phylogenetic trees. *Bioinformatics*. 30:2216–2218.
- 537 Pennell M.W., FitzJohn R.G., Cornwell W.K., Harmon L.J. 2015. Model Adequacy and
538 the Macroevolution of Angiosperm Functional Traits. *Am. Nat.* 186:E33-50.
- 539 Pennell M.W., Harmon L.J., Uyeda J.C. 2014b. Is there room for punctuated equilibrium
540 in macroevolution? *Trends Ecol. Evol.* 29:23–32.
- 541 Petchey O.L., Gaston K.J. 2006. Functional diversity: back to basics and looking
542 forward. *Ecol. Lett.* 9:741–758.
- 543 Pollock L.J., Rosauer D.F., Thornhill A.H., Kujala H., Crisp M.D., Miller J.T., McCarthy
544 M.A. 2015. Phylogenetic diversity meets conservation policy: small areas are key to
545 preserving eucalypt lineages. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*
546 370:20140007.
- 547 Popescu A.-A., Huber K.T., Paradis E. 2012. ape 3.0: New tools for distance-based
548 phylogenetics and evolutionary analysis in R. *Bioinformatics*. 28:1536–1537.
- 549 Rabosky D.L. 2010. Extinction rates should not be estimated from molecular
550 phylogenies. *Evolution (N. Y.)*. 64:1816–1824.
- 551 Rao R.C. 1982. Diversity and dissimilarity coefficients: a unified approach. *Theor. Popul.*
552 *Biol.* 21:24–43.
- 553 Revell L.J. 2012. phytools: an R package for phylogenetic comparative biology (and
554 other things). *Methods Ecol. Evol.* 3:217–223.
- 555 Ricotta C. 2005. A note on functional diversity measures. *Basic Appl. Ecol.* 6:479–486.
- 556 Robert C.P. 2007. *The Bayesian choice : from decision-theoretic foundations to*
557 *computational implementation*. Springer.
- 558 Rodrigues A.S.L., Gaston K.J. 2002. Maximising phylogenetic diversity in the selection
559 of networks of conservation areas. *Biol. Conserv.* 105:103–111.
- 560 Rosauer D.F., Mooers A.Ø. 2013. Nurturing the use of evolutionary diversity in nature
561 conservation. *Trends Ecol. Evol.* 28:322–3.
- 562 Schleuter D., Daufresne M., Massol F., Argillier C. 2010. A user’s guide to functional
563 diversity indices. *Ecol. Monogr.* 80:469–484.
- 564 Slater G.J., Harmon L.J., Wegmann D., Joyce P., Revell L.J., Alfaro M.E. 2012. Fitting
565 models of continuous trait evolution to incompletely sampled comparative data

- 566 using approximate bayesian computation. *Evolution* (N. Y). 66:752–762.
- 567 Srivastava D.S., Vellend M. 2005. Biodiversity-Ecosystem Function Research: Is It
568 Relevant to Conservation? *Annu. Rev. Ecol. Evol. Syst.* 36:267–294.
- 569 Sukumaran J., Economo E.P., Lacey Knowles L. 2016. Machine Learning Biogeographic
570 Processes from Biotic Patterns: A New Trait-Dependent Dispersal and
571 Diversification Model with Model Choice By Simulation-Trained Discriminant
572 Analysis. *Syst. Biol.* 65:525–545.
- 573 Thuiller W., Maiorano L., Mazel F., Guilhaumon F., Ficetola G.F., Lavergne S., Renaud
574 J., Roquet C., Mouillot D. 2015. Conserving the functional and phylogenetic trees of
575 life of European tetrapods. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 370:20140005.
- 576 Tucker C.M., Cadotte M.W., Carvalho S.B., Davies T.J., Ferrier S., Fritz S.A., Grenyer
577 R., Helmus M.R., Jin L.S., Mooers A.Ø., Pavoine S., Purschke O., Redding D.W.,
578 Rosauer D.F., Winter M., Mazel F. 2016. A guide to phylogenetic metrics for
579 conservation, community ecology and macroecology. *Biol. Rev. Camb. Philos. Soc.*
- 580 Uyeda J.C., Harmon L.J. 2014. A Novel Bayesian Method for Inferring and Interpreting
581 the Dynamics of Adaptive Landscapes from Phylogenetic Comparative Data. *Syst.*
582 *Biol.* 63:902–918.
- 583 Vane-Wright R.I., Humphries C.J., Williams P.H. 1991. What to protect?—Systematics
584 and the agony of choice. *Biol. Conserv.* 55:235–254.
- 585 Villéger S., Mason N., Mouillot D. 2008. New multidimensional functional diversity
586 indices for a multifaceted framework in functional ecology. *Ecology.* 89:2290–2301.
- 587 Winter M., Devictor V., Schweiger O. 2013. Phylogenetic diversity and nature
588 conservation: where are we? *Trends Ecol. Evol.* 28:199–204.
- 589
590

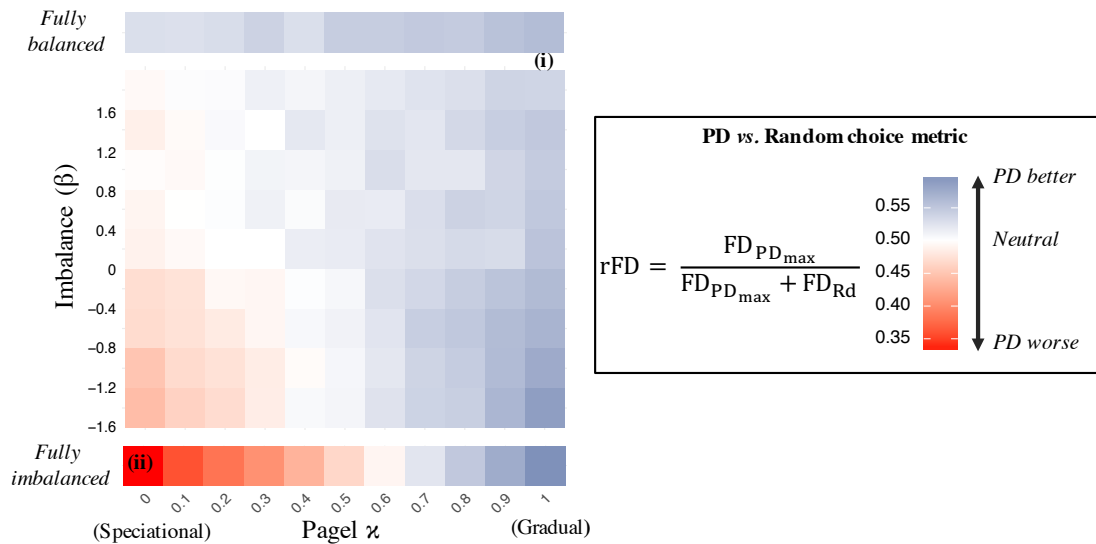
591

Type	Evolutionary Model			Type of Trees							
	Alpha	Beta	# traits	Yule Tree				Coalescent Tree			
				rFD		Phylo. Signal		rFD		Phylo. Signal	
				Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
BM	0	0	1	0.53	0.09	0.98	0.43	0.54	0.09	1.04	0.93
BM	0	0	2	0.55	0.13	1	0.31	0.6	0.12	1.01	0.64
BM	0	0	4	0.61	0.19	1	0.21	0.74	0.17	1.01	0.45
Markov			1	0.53	0.08			0.54	0.08		
Markov			2	0.53	0.07			0.56	0.08		
Markov			4	0.53	0.05			0.57	0.07		
EB	0	-5	1	0.55	0.09	4.32	2.51	0.56	0.11	9.18	7.36
EB	0	-1	1	0.53	0.09	1.34	0.62	0.55	0.09	1.52	1.33
EB	0	-5	2	0.62	0.13	4.33	1.83	0.66	0.15	9.32	5.55
EB	0	-1	2	0.57	0.12	1.33	0.43	0.61	0.13	1.52	1.01
EB	0	-5	4	0.77	0.16	4.23	1.36	0.82	0.16	9.23	4.17
EB	0	-1	4	0.65	0.17	1.34	0.35	0.74	0.17	1.49	0.67
OU	1.4	0	1	0.51	0.09	0.54	0.16	0.54	0.09	0.44	0.3
OU	7	0	1	0.51	0.1	0.24	0.06	0.52	0.09	0.14	0.06
OU	1.4	0	2	0.53	0.13	0.53	0.12	0.59	0.13	0.45	0.22
OU	7	0	2	0.51	0.13	0.24	0.04	0.55	0.13	0.14	0.04
OU	1.4	0	4	0.56	0.2	0.54	0.08	0.7	0.17	0.45	0.16
OU	7	0	4	0.51	0.2	0.23	0.03	0.59	0.19	0.14	0.03

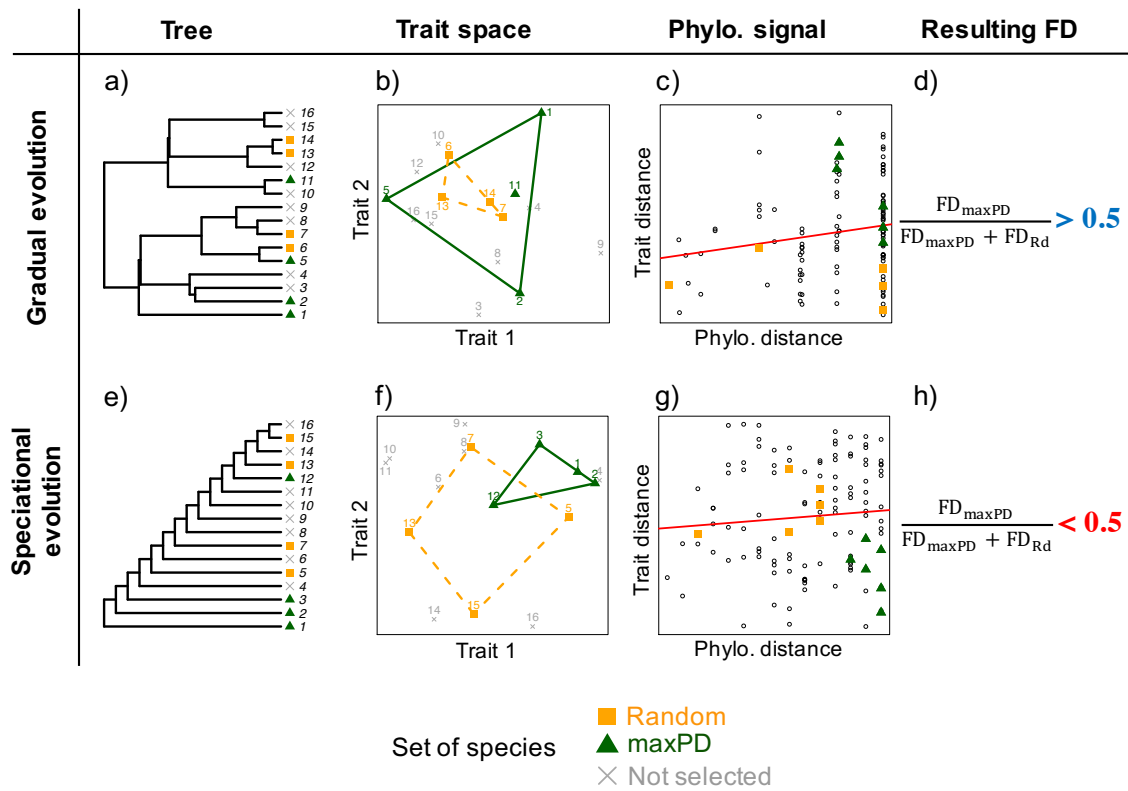
592

593 **Table 1. For common trait macroevolution models, sets of species that maximize PD**
594 **always harbour, on average, at least as much FD as random sets of species of the**
595 **same size.** The table presents, for each combination of macroevolutionary models
596 (column 1), specific set of parameters (column 2-3, the transition rate for the Markov
597 model is 1, see also methods) and number of independent traits (column 4), a measure of
598 the relative amount of FD (rFD) between maxPD and random sets of species for pure
599 birth Yule trees (column 5-6) and coalescent trees (column 9-10). These results
600 correspond to a tree of 64 species from which 8 are selected either at random or to
601 maximize PD (other combinations of these parameters are presented in Supp. Tables).
602 The comparison of FD (as captured by the convex hull measure) between the two sets of
603 species is quantified with the following metric: $rFD = FD_{maxPD} / (FD_{Random} + FD_{maxPD})$. A
604 value $<.5$ means PD is doing worse than random, a value $>.5$ means PD is doing better
605 than random and a value of $.5$ means PD is doing the same as random. The phylogenetic
606 signal for Yule trees (column 7-8) and coalescent trees (column 11-12) is measured with

607 the Blomberg K (for multiple traits, the mean across traits is given). All statistics are
 608 based on 1000 simulations in each case.
 609
 610



611
 612 **Figure 1. Under a speciation model trait evolution on imbalanced trees, sets of**
 613 **species that maximize PD harbour less FD than random sets of species of the same**
 614 **size.** The figure represents rFD (the relative amount of FD captured by the convex hull
 615 measure between the maxPD set and random sets of species) as a function of tree
 616 imbalance (as measured by β , Y-axis) and the degree of speciation vs. gradual
 617 evolution (as measured by Pagel κ , X-axis). The color of each grid cell reflects the mean
 618 value of the metric over 100 trait simulations on 10 different trees (for a total of 1000
 619 simulations) or, in the case of fully balanced and fully imbalanced trees, 1000 simulations
 620 on one single tree. Results are based on sets of 8 species out of 64 (tree size) and two
 621 traits. The two specific positions ‘i’ and ‘ii’ drawn on the figure refers to the parameter
 622 space position of the examples presented in figure 2, panels a-d and e-h, respectively. The
 623 tree presented in Figure 2a (corresponding to the position marked by ‘i’ in the present
 624 figure) has an imbalance of $\beta=3.5$.



625

626 **Figure 2. Examples of cases where the set of species that maximize PD harbours**
 627 **more (a-d) or less (e-h) FD than a random set of species.** For each example, the
 628 original phylogenetic tree (panels a and d), the position of species in trait space and their
 629 corresponding convex hull (panels b and f), the relationship between phylogenetic and
 630 trait distances (panels c and g) and the corresponding relative amount of FD between
 631 PD_{\max} set and random sets are given (panels d and h). Example (a-d) corresponds to a BM
 632 model on a relatively balanced tree while example (e-h) corresponds to a speciation
 633 model (Pagel $\kappa = 0$) on a fully imbalanced tree. Both examples are also reported in figure
 634 1, but note that here, for the purpose of simplicity, we used a tree with only 16
 635 species from which four species were selected.

636

637

SUPPLEMENTAL MATERIAL

638

639

for

640

641 *Conserving phylogenetic diversity can be a poor strategy for conserving functional diversity*

642

643 Florent Mazel^{1*}, Arne Mooers¹, Giulio Valentino Dalla Riva², and Matthew W. Pennell²

644

645 ¹ Department of Biological Sciences, Simon Fraser University, Burnaby, BC, Canada

646 ² Department of Zoology, Biodiversity Research Centre, University of British Columbia,

647 Vancouver, BC V6T 1Z4, Canada;

648

649 *flo.mazel@gmail.com

650

651

List of supplemental Materials

652

Supplemental Tables.

654

655 **Supplemental Table 1.** For common trait macroevolution models, sets of species that maximize
656 PD always harbour, on average, at least as much FD (Convex hull measure) as random sets of
657 species of the same size.

658 **Supplemental Table 2.** For common trait macroevolution models, sets of species that maximize
659 PD always harbour, on average, at least as much FD (Rao quadratic entropy measure) as random
660 sets of species of the same size.

661

Supplemental Figures.

663

664 **Figure S1.** Extension of our results to multiple traits and varying tree sizes and selected species
665 number using Convex Hull as a measure of FD.

666 **Figure S2.** Extension of our results to multiple traits and varying tree sizes and selected species
667 number using Rao quadratic entropy as a measure of FD.

668 **Figure S3.** Extension of our results to discrete trait evolution.

669 **Figure S4.** Variability of the phylogenetic signal as measured by the Bloomberg K.

670 **Figure S5.** Variability of the phylogenetic signal as measured by the Spearman correlation
 671 between trait and phylogenetic distances.
 672

Type	Evolutionary Model		Traits and Species			Type of Trees							
	α	r	# traits	# species	# selected species	Yule Tree				Coalescent Tree			
						FD ratio metric		Phylo. Signal		FD ratio metric		Phylo. Signal	
						Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd
			1	32	8	0.53	0.08	0.99	0.46	0.54	0.08	1	0.85
			1	64	8	0.53	0.09	0.98	0.43	0.54	0.09	1.04	0.93
			1	32	16	0.51	0.05	1.01	0.48	0.52	0.05	0.97	0.78
			1	64	16	0.52	0.06	1	0.41	0.53	0.05	1.02	0.96
			2	32	8	0.56	0.12	0.99	0.34	0.6	0.12	0.99	0.59
			2	64	8	0.55	0.13	1	0.31	0.6	0.12	1.01	0.64
			2	32	16	0.53	0.07	0.97	0.31	0.55	0.06	1	0.57
			2	64	16	0.54	0.08	1	0.28	0.56	0.08	0.98	0.61
			4	32	8	0.63	0.18	1	0.24	0.73	0.16	0.99	0.42
			4	64	8	0.61	0.19	1	0.21	0.74	0.17	1.01	0.45
			4	32	16	0.58	0.09	1	0.24	0.62	0.1	1	0.42
			4	64	16	0.59	0.11	1	0.21	0.66	0.11	0.97	0.41
		-5	1	32	8	0.56	0.09	3.42	1.91	0.55	0.1	6.56	4.7
		-5	1	64	8	0.55	0.09	4.32	2.51	0.56	0.11	9.18	7.36
		-5	1	32	16	0.52	0.05	3.37	1.86	0.52	0.06	6.6	4.69
		-5	1	64	16	0.54	0.06	4.28	2.56	0.53	0.07	8.86	7.37
		-1	1	32	8	0.53	0.08	1.3	0.65	0.54	0.08	1.44	1.25
		-1	1	64	8	0.53	0.09	1.34	0.62	0.55	0.09	1.52	1.33
		-1	1	32	16	0.52	0.05	1.31	0.68	0.52	0.04	1.42	1.29
		-1	1	64	16	0.52	0.05	1.34	0.64	0.53	0.06	1.54	1.4
		-5	2	32	8	0.62	0.12	3.35	1.35	0.63	0.14	6.57	3.58
		-5	2	64	8	0.62	0.13	4.33	1.83	0.66	0.15	9.32	5.55
		-5	2	32	16	0.56	0.08	3.22	1.25	0.56	0.09	6.69	3.54
		-5	2	64	16	0.58	0.08	4.26	1.84	0.58	0.1	9.33	5.58
		-1	2	32	8	0.57	0.12	1.29	0.45	0.6	0.12	1.44	0.94
		-1	2	64	8	0.57	0.12	1.33	0.43	0.61	0.13	1.52	1.01
		-1	2	32	16	0.54	0.07	1.28	0.44	0.55	0.07	1.4	0.9
		-1	2	64	16	0.55	0.08	1.32	0.45	0.57	0.08	1.5	0.95
		-5	4	32	8	0.77	0.16	3.36	1.1	0.79	0.17	6.45	2.61
		-5	4	64	8	0.77	0.16	4.23	1.36	0.82	0.16	9.23	4.17
		-5	4	32	16	0.65	0.11	3.35	1.08	0.65	0.12	6.24	2.52
		-5	4	64	16	0.69	0.12	4.4	1.54	0.69	0.13	9.42	4.09
		-1	4	32	8	0.67	0.17	1.28	0.33	0.73	0.17	1.42	0.62
		-1	4	64	8	0.65	0.17	1.34	0.35	0.74	0.17	1.49	0.67
		-1	4	32	16	0.6	0.09	1.28	0.33	0.62	0.1	1.41	0.64
		-1	4	64	16	0.62	0.11	1.33	0.31	0.67	0.11	1.44	0.66
	1.4	0	1	32	8	0.51	0.09	0.6	0.21	0.53	0.08	0.47	0.33
	1.4	0	1	64	8	0.51	0.09	0.54	0.16	0.54	0.09	0.44	0.3
	1.4	0	1	32	16	0.51	0.05	0.59	0.23	0.52	0.04	0.46	0.3
	1.4	0	1	64	16	0.51	0.06	0.54	0.16	0.52	0.05	0.45	0.3
	7	0	1	32	8	0.5	0.09	0.28	0.1	0.51	0.09	0.17	0.08
	7	0	1	64	8	0.51	0.1	0.24	0.06	0.52	0.09	0.14	0.06
	7	0	1	32	16	0.5	0.05	0.28	0.1	0.52	0.05	0.17	0.07
	7	0	1	64	16	0.5	0.07	0.24	0.06	0.51	0.06	0.14	0.06
	1.4	0	2	32	8	0.53	0.12	0.59	0.16	0.57	0.12	0.47	0.23
	1.4	0	2	64	8	0.53	0.13	0.53	0.12	0.59	0.13	0.45	0.22
	1.4	0	2	32	16	0.52	0.07	0.58	0.16	0.54	0.06	0.47	0.23
	1.4	0	2	64	16	0.52	0.08	0.54	0.12	0.56	0.08	0.43	0.2
	7	0	2	32	8	0.51	0.13	0.28	0.08	0.54	0.13	0.16	0.06
	7	0	2	64	8	0.51	0.13	0.24	0.04	0.55	0.13	0.14	0.04
	7	0	2	32	16	0.51	0.07	0.28	0.07	0.53	0.06	0.17	0.06
	7	0	2	64	16	0.51	0.09	0.24	0.05	0.54	0.08	0.14	0.04
	1.4	0	4	32	8	0.57	0.18	0.58	0.11	0.69	0.17	0.48	0.17
	1.4	0	4	64	8	0.56	0.2	0.54	0.08	0.7	0.17	0.45	0.16
	1.4	0	4	32	16	0.56	0.1	0.59	0.11	0.61	0.09	0.48	0.17
	1.4	0	4	64	16	0.55	0.11	0.55	0.09	0.64	0.11	0.44	0.15
	7	0	4	32	8	0.52	0.19	0.28	0.06	0.6	0.19	0.17	0.04
	7	0	4	64	8	0.51	0.2	0.23	0.03	0.59	0.19	0.14	0.03
	7	0	4	32	16	0.51	0.1	0.28	0.06	0.58	0.1	0.17	0.04
	7	0	4	64	16	0.52	0.12	0.23	0.03	0.59	0.11	0.14	0.03

673

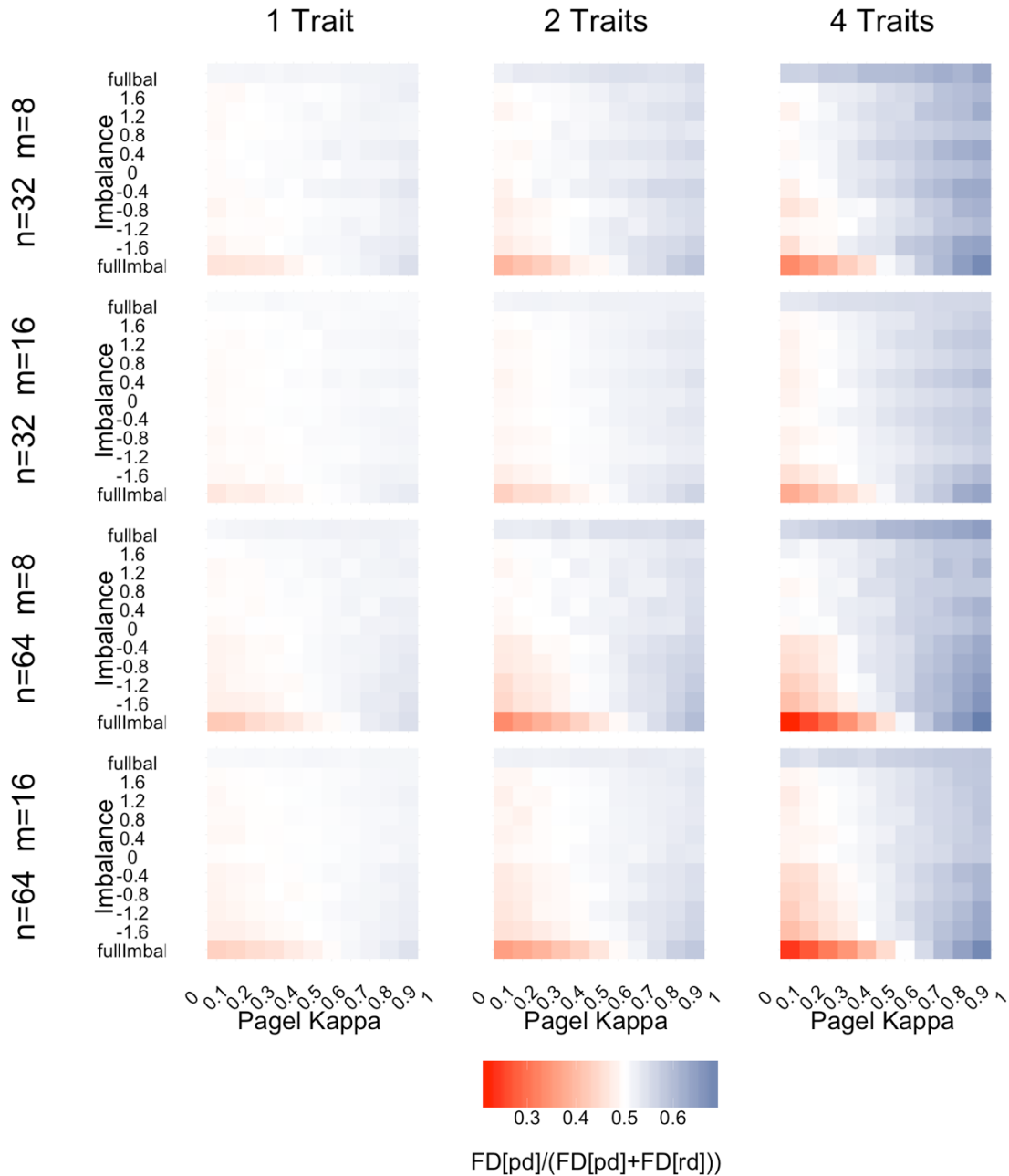
674 **Supplemental Table 1. For common trait macroevolution models, sets of species that**
 675 **maximize PD always harbour, on average, at least as much FD (Convex hull measure) as**
 676 **random sets of species of the same size.** The table presents, for each combination of
 677 macroevolutionary models (column 1), specific set of parameters (column 2-3), number of
 678 independent traits (column 4), tree size (column 5) and number of selected species (column 6) a
 679 measure of the relative amount of FD between PD_{max} and random sets of species for pure birth
 680 Yule trees (column 7-8) and coalescent trees (column 11-12). The comparison of FD (as measured

681 by the convex hull measure) between the two sets of species is quantified with the following
 682 metric: $FD_{\max PD} / (FD_{\text{Random}} + FD_{\max PD})$. A value $<.5$ means PD is doing worse than random, a value $>.5$
 683 means PD is doing better than random and a value of $.5$ means PD is doing the same as
 684 random. The phylogenetic signal for Yule trees (column 9-10) and coalescent trees (column 13-
 685 14) is measured with the Bloomberg K (for multiple traits, the mean across traits is given). All
 686 statistics are based on 1000 simulations in each case.
 687

Type	Evolutionary Model		Traits and Species				Type of Trees							
	α	r	# traits	# species	# selected species	Yule Tree				Coalescent Tree				
						FD ratio metric		Phylo. Signal		FD ratio metric		Phylo. Signal		
						Mean	Sd	Mean	Sd	Mean	Sd	Mean	Sd	
BM	0	0	1	32	8	0.54	0.15	0.99	0.46	0.57	0.15	1	0.85	
BM	0	0	1	64	8	0.55	0.16	0.98	0.43	0.57	0.15	1.04	0.93	
BM	0	0	1	32	16	0.52	0.08	1.01	0.48	0.52	0.09	0.97	0.78	
BM	0	0	1	64	16	0.52	0.1	1	0.41	0.54	0.1	1.02	0.96	
BM	0	0	2	32	8	0.54	0.11	0.99	0.34	0.57	0.11	0.99	0.59	
BM	0	0	2	64	8	0.54	0.12	1	0.31	0.56	0.12	1.01	0.64	
BM	0	0	2	32	16	0.52	0.06	0.97	0.31	0.52	0.06	1	0.57	
BM	0	0	2	64	16	0.53	0.07	1	0.28	0.54	0.08	0.98	0.61	
BM	0	0	4	32	8	0.54	0.08	1	0.24	0.56	0.08	0.99	0.42	
BM	0	0	4	64	8	0.54	0.08	1	0.21	0.57	0.09	1.01	0.45	
BM	0	0	4	32	16	0.52	0.04	1	0.24	0.53	0.05	1	0.42	
BM	0	0	4	64	16	0.52	0.05	1	0.21	0.54	0.06	0.97	0.41	
EB	0	-5	1	32	8	0.59	0.15	3.42	1.91	0.57	0.16	6.56	4.7	
EB	0	-5	1	64	8	0.58	0.16	4.32	2.51	0.59	0.17	9.18	7.36	
EB	0	-5	1	32	16	0.53	0.09	3.37	1.86	0.53	0.1	6.6	4.69	
EB	0	-5	1	64	16	0.55	0.11	4.28	2.56	0.55	0.12	8.86	7.37	
EB	0	-1	1	32	8	0.55	0.14	1.3	0.65	0.57	0.14	1.44	1.25	
EB	0	-1	1	64	8	0.54	0.16	1.34	0.62	0.57	0.15	1.52	1.33	
EB	0	-1	1	32	16	0.52	0.08	1.31	0.68	0.52	0.09	1.42	1.29	
EB	0	-1	1	64	16	0.52	0.1	1.34	0.64	0.53	0.11	1.54	1.4	
EB	0	-5	2	32	8	0.58	0.12	3.35	1.35	0.58	0.14	6.57	3.58	
EB	0	-5	2	64	8	0.58	0.12	4.33	1.83	0.6	0.15	9.32	5.55	
EB	0	-5	2	32	16	0.54	0.07	3.22	1.25	0.53	0.09	6.69	3.54	
EB	0	-5	2	64	16	0.54	0.08	4.26	1.84	0.54	0.11	9.33	5.58	
EB	0	-1	2	32	8	0.55	0.11	1.29	0.45	0.57	0.12	1.44	0.94	
EB	0	-1	2	64	8	0.55	0.11	1.33	0.43	0.57	0.12	1.52	1.01	
EB	0	-1	2	32	16	0.52	0.06	1.28	0.44	0.53	0.07	1.4	0.9	
EB	0	-1	2	64	16	0.53	0.07	1.32	0.45	0.54	0.08	1.5	0.95	
EB	0	-5	4	32	8	0.58	0.1	3.36	1.1	0.58	0.13	6.45	2.61	
EB	0	-5	4	64	8	0.59	0.1	4.23	1.36	0.59	0.14	9.23	4.17	
EB	0	-5	4	32	16	0.53	0.06	3.35	1.08	0.53	0.08	6.24	2.52	
EB	0	-5	4	64	16	0.55	0.07	4.4	1.54	0.55	0.1	9.42	4.09	
EB	0	-1	4	32	8	0.55	0.08	1.28	0.33	0.57	0.1	1.42	0.62	
EB	0	-1	4	64	8	0.55	0.08	1.34	0.35	0.57	0.1	1.49	0.67	
EB	0	-1	4	32	16	0.52	0.05	1.28	0.33	0.53	0.06	1.41	0.64	
EB	0	-1	4	64	16	0.53	0.05	1.33	0.31	0.54	0.07	1.44	0.66	
OU	1.4	0	1	32	8	0.52	0.15	0.6	0.21	0.56	0.14	0.47	0.33	
OU	1.4	0	1	64	8	0.52	0.16	0.54	0.16	0.56	0.16	0.44	0.3	
OU	1.4	0	1	32	16	0.52	0.09	0.59	0.23	0.53	0.08	0.46	0.3	
OU	1.4	0	1	64	16	0.52	0.11	0.54	0.16	0.53	0.1	0.45	0.3	
OU	7	0	1	32	8	0.51	0.16	0.28	0.1	0.52	0.16	0.17	0.08	
OU	7	0	1	64	8	0.51	0.17	0.24	0.06	0.53	0.16	0.14	0.06	
OU	7	0	1	32	16	0.5	0.09	0.28	0.1	0.52	0.09	0.17	0.07	
OU	7	0	1	64	16	0.5	0.11	0.24	0.06	0.52	0.1	0.14	0.06	
OU	1.4	0	2	32	8	0.52	0.11	0.59	0.16	0.55	0.1	0.47	0.23	
OU	1.4	0	2	64	8	0.52	0.12	0.53	0.12	0.56	0.11	0.45	0.22	
OU	1.4	0	2	32	16	0.52	0.06	0.58	0.16	0.52	0.06	0.47	0.23	
OU	1.4	0	2	64	16	0.51	0.07	0.54	0.12	0.53	0.07	0.43	0.2	
OU	7	0	2	32	8	0.51	0.11	0.28	0.08	0.53	0.11	0.16	0.06	
OU	7	0	2	64	8	0.51	0.12	0.24	0.04	0.53	0.12	0.14	0.04	
OU	7	0	2	32	16	0.51	0.06	0.28	0.07	0.52	0.06	0.17	0.06	
OU	7	0	2	64	16	0.5	0.08	0.24	0.05	0.52	0.07	0.14	0.04	
OU	1.4	0	4	32	8	0.52	0.08	0.58	0.11	0.55	0.08	0.48	0.17	
OU	1.4	0	4	64	8	0.52	0.09	0.54	0.08	0.56	0.08	0.45	0.16	
OU	1.4	0	4	32	16	0.51	0.04	0.59	0.11	0.52	0.04	0.48	0.17	
OU	1.4	0	4	64	16	0.51	0.05	0.55	0.09	0.53	0.06	0.44	0.15	
OU	7	0	4	32	8	0.51	0.08	0.28	0.06	0.53	0.08	0.17	0.04	
OU	7	0	4	64	8	0.5	0.08	0.23	0.03	0.53	0.08	0.14	0.03	
OU	7	0	4	32	16	0.5	0.04	0.28	0.06	0.52	0.04	0.17	0.04	
OU	7	0	4	64	16	0.51	0.05	0.23	0.03	0.52	0.05	0.14	0.03	

688
 689 **Supplemental Table 2. For common trait macroevolution models, sets of species that**
 690 **maximize PD always harbour, on average, at least as much FD (Rao quadratic entropy**
 691 **measure) as random sets of species of the same size.** The table presents, for each combination
 692 of macroevolutionary models (column 1), specific set of parameters (column 2-3), number of
 693 independent traits (column 4), tree size (column 5) and number of selected species (column 6) a
 694 measure of the relative amount of FD between PD_{\max} and random sets of species for pure birth

695 Yule trees (column 7-8) and coalescent trees (column 11-12). The comparison of FD (as measured
696 by the Rao quadratic entropy) between the two sets of species is quantified with the following
697 metric: $FD_{\max PD} / (FD_{\text{random}} + FD_{\max PD})$. A value $<.5$ means PD is doing worse than random, a value $>.5$
698 means PD is doing better than random and a value of $.5$ means PD is doing the same as
699 random. The phylogenetic signal for Yule trees (column 9-10) and coalescent trees (column 13-
700 14) is measured with the Bloomberg K (for multiple traits, the mean across traits is given). All
701 statistics are based on 1000 simulations in each case.
702

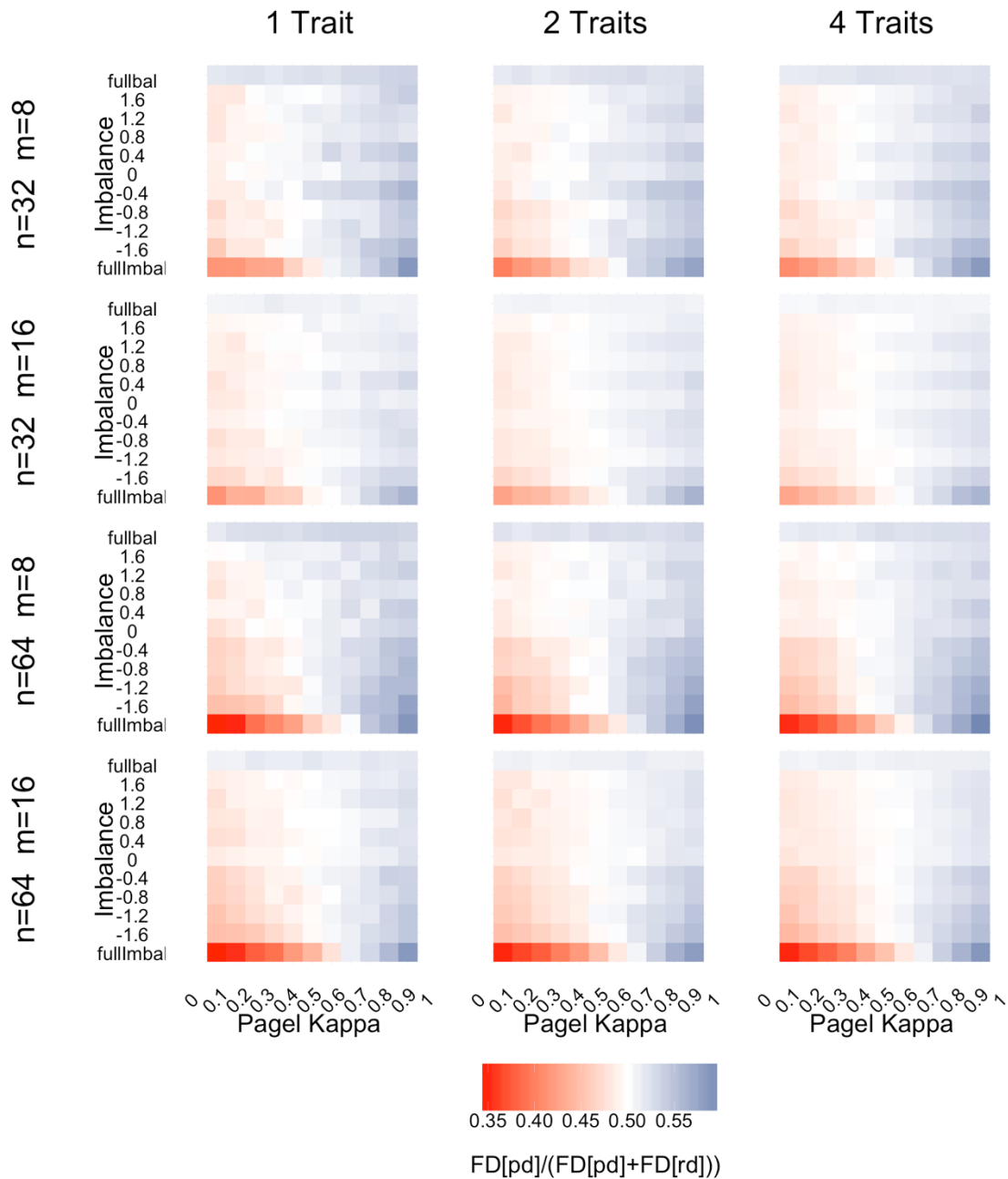


703

704 **Figure S1. Extension of our results to multiple traits and varying tree sizes and selected**
 705 **species number using Convex Hull as a measure of FD.** The figure present the variation of a
 706 measure of the relative amount of FD (as measured by the convex Hull measure) between PD-
 707 maximized and random set of species (see legend) in function of tree imbalance (as measured by
 708 β , Y-axis, “fullbal” refers to fully balanced tree and “fullImbal” refers to fully imbalanced
 709 tree) and the degree of speciation vs. gradual evolution (as measured by Pagel κ , X-axis). The
 710 color of each grid cell reflects the mean value of the metric over 100 trait simulations on 10

711 different trees (for a total of 1000 simulations) or, in the case of fully balanced and fully
712 imbalanced trees, 1000 simulations on one single tree. Each panel corresponds to a different set
713 of parameters (tree size (n), selected number of species (m) and number of traits).

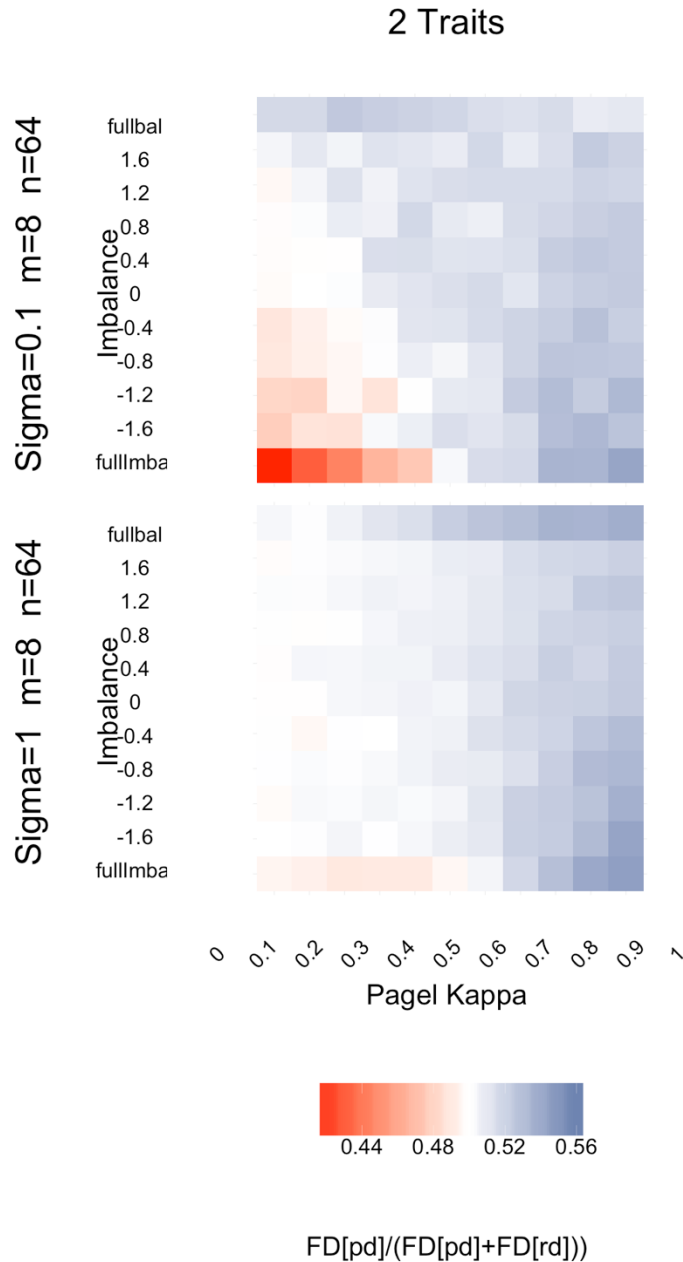
714



715

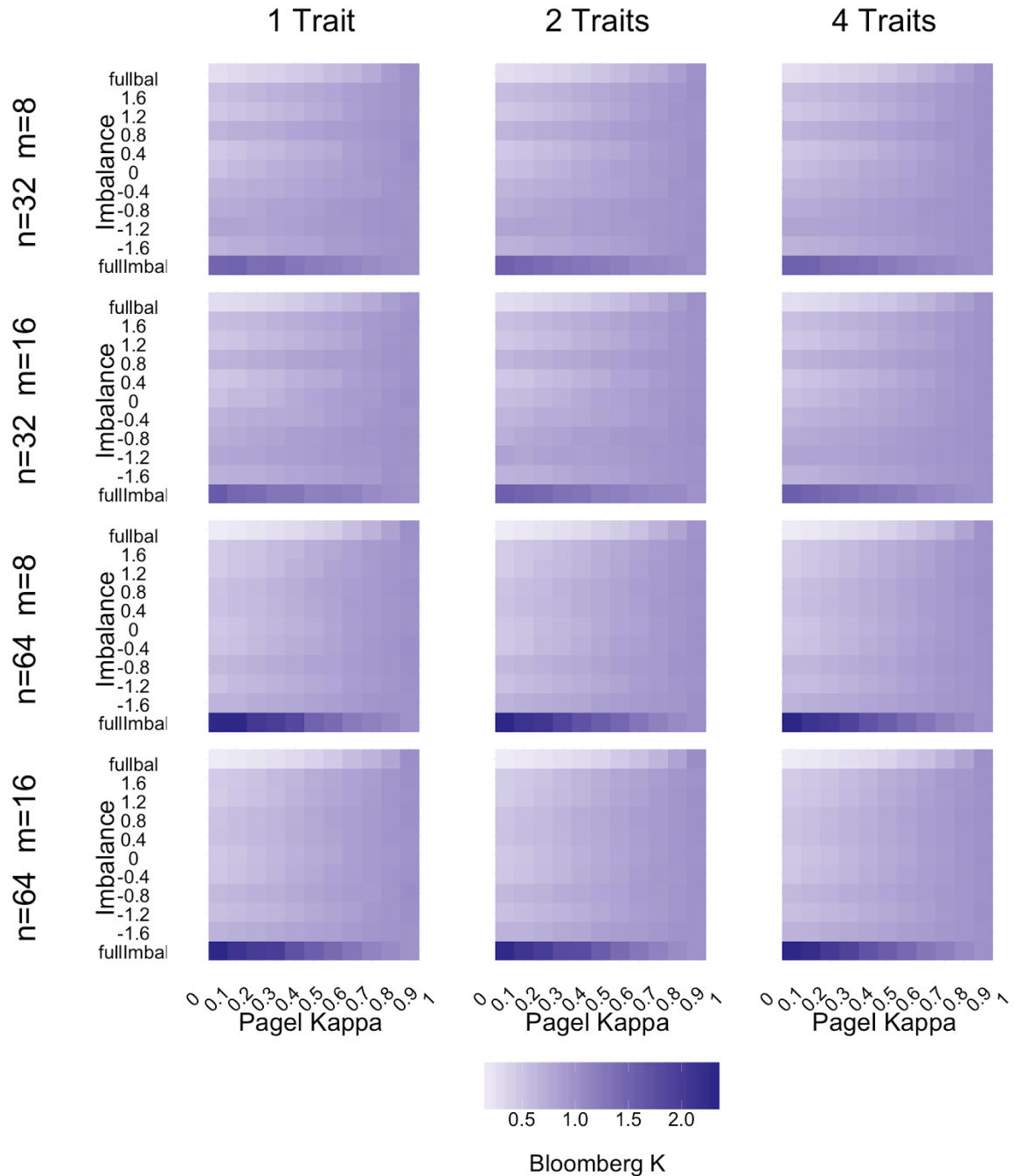
716 **Figure S2. Extension of our results to multiple traits and varying tree sizes and selected**
717 **species number using Rao quadratic entropy as a measure of FD.** The figure presents the
718 variation of a measure of the relative amount of FD (as measured by Rao quadratic entropy)

719 between PD-maximized and random set of species (see legend) in function of tree imbalance (as
720 measured by β , Y-axis, “fullbal” refers to fully balanced tree and “fullImbal” refers to fully
721 imbalanced tree) and the degree of speciation vs. gradual evolution (as measured by Pagel κ , X-
722 axis). The color of each grid cell reflects the mean value of the metric over 100 trait simulations
723 on 10 different trees (for a total of 1000 simulations) or, in the case of fully balanced and fully
724 imbalanced trees, 1000 simulations on one single tree. Each panel corresponds to a different set
725 of parameters (tree size (n), selected number of species (m) and number of traits).



727

728 **Figure S3. Extension of our results to discrete trait evolution.** The figure presents the variation
729 of a measure of the relative amount of FD (as measured by the number of character state
730 combinations) between PD-maximized and random set of species (see legend) in function of tree
731 imbalance (as measured by β , Y-axis, “fullbal” refers to fully balanced tree and “fullImbal”
732 refers to fully imbalanced tree) and the degree of speciation vs. gradual evolution (as measured
733 by Pagel κ , X-axis). The color of each grid cell reflects the mean value of the metric over 100
734 trait simulations on 10 different trees (for a total of 1000 simulations) or, in the case of fully
735 balanced and fully imbalanced trees, 1000 simulations on one single tree. Results are based on
736 sets of 8 species out of 64 (tree size). The different panels correspond to different values of the
737 transition rate parameter of the Markov model.

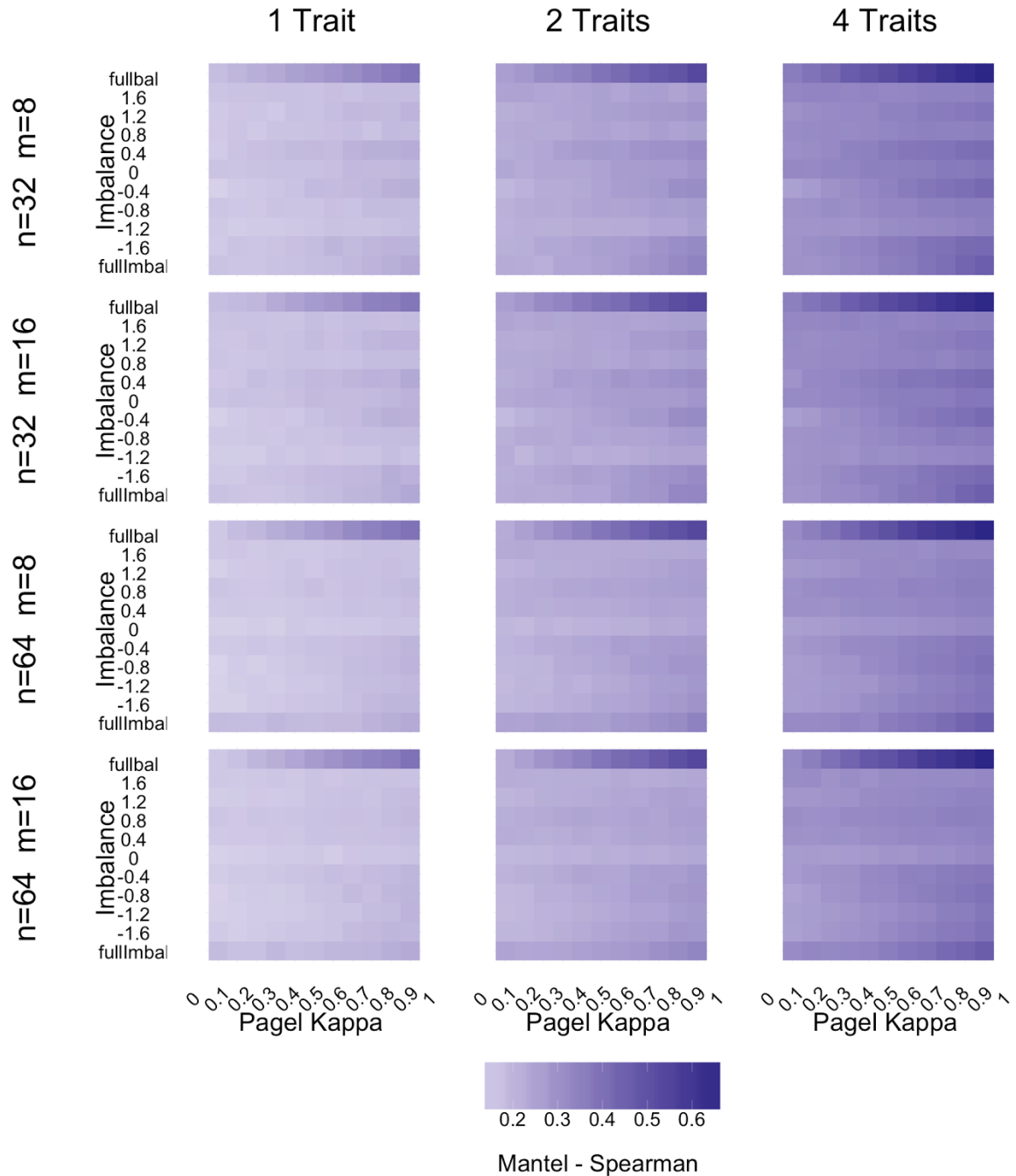


738

739

740 **Figure S4. Variability of the phylogenetic signal as measured by the Bloomberg K.** The
741 figure presents the variation of phylogenetic signal across the parameter space presented on figure
742 1 of the main text. Mean Bloomberg K statistic (see legend) are presented in function of tree
743 imbalance (as measured by β , Y-axis, “fullbal” refers to fully balanced tree and “fullImbal”
744 refers to fully imbalanced tree) and the degree of speciation vs. gradual evolution (as measured

745 by Pagel κ , X-axis). Each panel corresponds to a different set of parameters (tree size (n),
746 selected number of species (m) and number of traits). The color of each grid cell reflects the mean
747 value of the metric over 100 trait simulations on 10 different trees (for a total of 1000
748 simulations) or, in the case of fully balanced and fully imbalanced trees, 1000 simulations on one
749 single tree.



750

751 **Figure S5. Variability of the phylogenetic signal as measured by the Spearman correlation**
752 **between trait and phylogenetic distances.** The figure presents the variation of phylogenetic

753 signal across the parameter space presented on figure 1 of the main text. Mean mantel
754 (spearman) statistic (see legend) are presented in function of tree imbalance (as measured by β ,
755 Y-axis, “fullbal” refers to fully balanced tree and “fullImbal” refers to fully imbalanced tree)
756 and the degree of speciation vs. gradual evolution (as measured by Pagel κ , X-axis). Each panel
757 corresponds to a different set of parameters (tree size (n), selected number of species (m) and
758 number of traits). The color of each grid cell reflects the mean value of the metric over 100 trait
759 simulations on 10 different trees (for a total of 1000 simulations) or, in the case of fully balanced
760 and fully imbalanced trees, 1000 simulations on one single tree.
761